Neural networks enhanced adaptive admittance control of optimized robot-environment interaction

Article (Accepted Version)

Neural Networks Enhanced Adaptive Admittance Control of Optimized Robot-Environment Interaction

Chenguang Yang, Guangzhu Peng, Yanan Li, Rongxin Cui, Long Cheng, and Zhijun Li

***Abstract*—In this paper, an admittance adaptation method for robots to interact with unknown environment is developed. The environment to be interact with is modelled as a linear system in the state-space form. In the presence of the unknown environment dynamics, an observer in robot joint space is employed to estimate the interaction torque. Admittance control is adopted to regulate the dynamic behavior at the interaction point when the robot interacts with unknown environment. An adaptive neural controller using radial basis function is employed to guarantee trajectory tracking. A cost function is defined to achieve the interaction performance of torque regulation and trajectory tracking, which is minimised by adaptation of the admittance model. Simulation studies on a robot manipulator are carried out to verify the effectiveness of the proposed method.**

***Index Terms*—optimal adaptive control; robot−environment interaction; observer; neural networks (NNs); admittance control**

I. INTRODUCTION

WITH the development of robot technology, robots have been widely used in various fields such as education, industry and entertainment, etc. In these applications, robots are required to interact with external environment [1]–[3]. Therefore, robots interacting with the environment has received great attention and much effort has been made on this topic. Although it has been investigated for more than decades, there are still many open problems not solved,, due to the high expectation of robots in more general scenarios and the complex environment in which robots are working. In order to achieve a compliant behavior, there are three approaches that are widely applied: admittance control, hybrid position/force control and impedance control.

The concept of impedence control introduced by Hogan has been a classical control method in robotics [4]. The aim of impedence control is to develop a relationship between the interaction force and the position of the robot. The core idea of impedence control is that the controller should modulate the mechanical impedence, which is a mapping from generalized velocities to generalized force. This control approach appears to be feasible and robust [5]. Another approach is admittance control, which was introduced by Mason [6]. In a generalized admittance control system, with the measurements of environment force and a desired admittance model, a virtual desired trajectory is obtained and tracked. Then, the compliant behavior is realized by trajectory adaptation. Traditional control method of a robot manipulator is model-based control, which usually has a good control performance [7]. However, this method heavily depends on the accuracy of a robot model which cannot be guaranteed in many cases. Therefore, adaptive control methods have been widely studied and applied to practical systems [8]. These methods can approximate uncertainties of a system by using tools such as neural network (NN), wavelet network, and fuzzy logic system, etc [9]–[11]. Another key element in admittance control system is the force sensor. Force sensors are regarded as a media for communication between a robot and environment. However, force sensors equipped on the manipulators may cause inconvenience and are usually costly. Due to these reasons, sensorless control schemes have been received great attention. There are two main methods for estimating the external force: disturbance observer approach and force observer approach based on knowledge of motor torques. In [12], the disturbance observer approach with knowledge of joint angle has been analyzed. In [13], a force observer for collision detection based on the generalized momentum has been introduced. In [14], an collision detection method is first developed for rigid robot arms and other robots with elastic joints.

Under impedence/admittance control, robots are governed to be compliant to interaction force exerted by the environment. If the environment is passive, a passive impedence/admittance model is imposed to the robot for safety. However, obtaining desired impedence/admittance model is not a easy work due to the complexity of the environmental dynamics. Moreover, a fixed impedence/admittance model could not be applied to all situations. To solve this problem, iterative learning has been widely studied for robots to adapt to unknown environments. This approach aims to introduce human learning skills to robots. It has been generally acknowledged that such an ability of improvng performance by repeating a task is an important control strategy and has been widely studied. In [15], an associative search network (ASN) learning scheme is presented for learning control parameters for robot to complete a wall-following task. In [16], neural networks based method is applied to regulate impedance parameters of the end-effector of a robot. However, the disadvantage of the learning method is that it requires a robot to repeat operations to learn the desired impedance parameters which may cause inconvenience in many situations. Therefore, the impedence adaptation method has been widely studied [17]. In [18], strategies of switching

This work was partially supported by National Nature Science Foundation (NSFC) under Grant 61473120, Science and Technology Planning Project of Guangzhou 201607010006 and the Fundamental Research Funds for the Central Universities under Grant 2015ZM065.

C. Yang, G. Peng and Z. Li are with Key Laboratory of Autonomous Systems and Networked Control, College of Automation Science and Engineering, South China University of Technology, Guangzhou, 510640 China. Emails: cyang@ieee.org; gz.peng@qq.com; zjli@ieee.org.

Yanan Li is with Department of Engineering and Design, University of Sussex, Brighton, BN1 9RH, UK. Email: yl557@sussex.ac.uk.

Rongxin Cui is with School of Marine Science and Technology, Northwestern Polytechnical University, China. Email: r.cui@nwpu.edu.cn.

Long Cheng is with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. Email: long.cheng@ia.ac.cn.

among different values of impedance parameters are developed for dissipating the energy of the system. In [19], the impedance adaptation is investigated for robots to interact with unknown environments.

The control objective of interaction control is to achieve force regulation and trajectory tracking. Thus, optimisation should be taken into consideration, since it is the compromise of these two objectives. There has been much research effort in literatures. The well known linear quadratic regulator (LQR) is widely acknowledged as an important solution of optimal control, which is concerned with operating a dynamic system at a minimal cost. In [20], the LQR is used to determine desired impedance parameters with the environmental dynamics known. In [21], a target impedance is adjusted by online solutions of the defined LQR problem based on environment stiffness and damping. Better than the fixed impedance parameters obtained from LQR technique, the algorithm shows greater adaptability for a wide range of environments. However, the dynamics of the environment is also assumed to be known in this paper above. As presented in [22], the solution of a Riccati equation could be difficult to find with the unknown dynamics of a environment. Therefore, when the dynamics of a environment is unknown, approaches proposed above may not be used. To copy with this problem, adaptive dynamic programming (ADP) has received much attention and been widely studied [23]–[27]. ADP is a very useful tool in solving optimization and optimal control problems. Based on the idea of ADP, a control action is modified based on the feedback information of a environment. There are many ADP approaches such as heuristic dynamic programming (HDP), Q-learning and dual-heuristic programming (DHP). The advantage of ADP is that only partial information of the system under control needs to be known. In [28], optimal impedance parameters are updated by employing a recursive least-square filter-based episodic natural actor-critic algorithm. In [29], the reinforcement learning (RL) algorithm is adopted to accomplish variable impedance control. However, in many situations, the process of learning is still needed to obtain parameters of a impedence/admittance model [30]. The optimal control method with unknown environment proposed in [31] is to be applied and developed. The environment model is considered to be a damping-stiffness model, which is a linear system with unknown dynamics.

Based on the above discussion, we propose a method to adjust admittance parameters subject to an unknown environment. First, a cost function is defined to describe the interaction performance including the trajectory tracking error and interaction torque. Second, an environment with unknown dynamics is taken into consideration and its model is described as a linear system in the state-space form. An observer based on the generalized momentum approach is used to estimate the interaction torque in joint space. As the environmental dynamics is unknown, admittance adaptation is proposed to adjust admittance parameters. As a result, the target admittance model which guarantees the optimal interaction behavior is obtained.

The rest of this paper is organized as follows. Section II describes the dynamics of a robot and control objective as well as the unknown environmental dynamics. Section III introduces the methodology including the estimation of the external torque and the admittance adaptation. In Section IV, the proposed method is verified through the simulations and the conclusion is drawn in Section V.

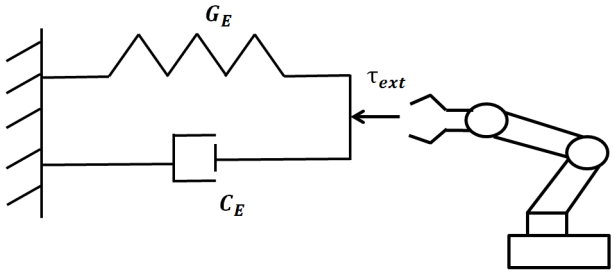

Fig. 1. The model of damping-stiffness environment

II. PRELIMINARIES

A. System Description

In this system, we consider that a robot arm is interacting with an environment. The kinematics of a robot can be expressed by

$$x(t) = \kappa(q) \tag{1}$$

where $\kappa(\cdot)$, $x(t)$, $q \in \mathbb{R}^n$ and n denote forward kinematics function, positions/orientations in the Cartesian space, joint vectors in joint space, and number of degrees of freedom (DOF), respectively. Differentiating (1) with respect to time results in

$$\dot{x}(t) = J(q)\dot{q} \tag{2}$$

where $J(q) \in \mathbb{R}^{n\times n}$ is the Jacobian matrix which is assumed to be non-singular in a finite work space. Further differentiating (2) with respect to time results in

$$\ddot{x}(t) = \dot{J}(q)\dot{q} + J(q)\ddot{q} \tag{3}$$

The dynamics of a robot arm in joint space can be given by

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) + \tau_{ext} = \tau \tag{4}$$

where $M(q) \in \mathbb{R}^{n\times n}$ is the inertia matrix in joint space; $C(q,\dot{q}) \in \mathbb{R}^{n\times n}$ is the Coriolis and centripetal coupling matrix; $G(q) \in \mathbb{R}^{n\times 1}$ is the gravity loading, $\dot{q}$ and $\ddot{q}$ are respectively the vector of generalized joint coordinates: velocities and accelerations; $\tau \in \mathbb{R}^n$ is the control input; $\tau_{ext} \in \mathbb{R}^n$ denotes the vector of joint torque exerted by the environment.

Property 1: Matrix $M(q)$ is symmetric and positive definite [32].

Property 2: Matrix $2C(q,\dot{q}) - \dot{M}(q)$ is a skew-symmetric matrix [32].

Now, let us consider the modelling of the environment. In this paper, a damping-stiffness environment model is considered, as shown in Fig. 1:

$$C_E\dot{q} + G_E q = -\tau_{ext} \tag{5}$$

where C_E and G_E are unknown damping and stiffness matrices of the environment.

Remark 1: Without loss of generality, the spring-damper system is usually used to describe the model of an environment. Compared to the model (5), a simple second-order equation is used as the model for an environment. In the model, mass, damping and stiffness are considered: $M_E\ddot{q} + C_E\dot{q} + G_E q = -\tau_{ext}$. In general, a large range of

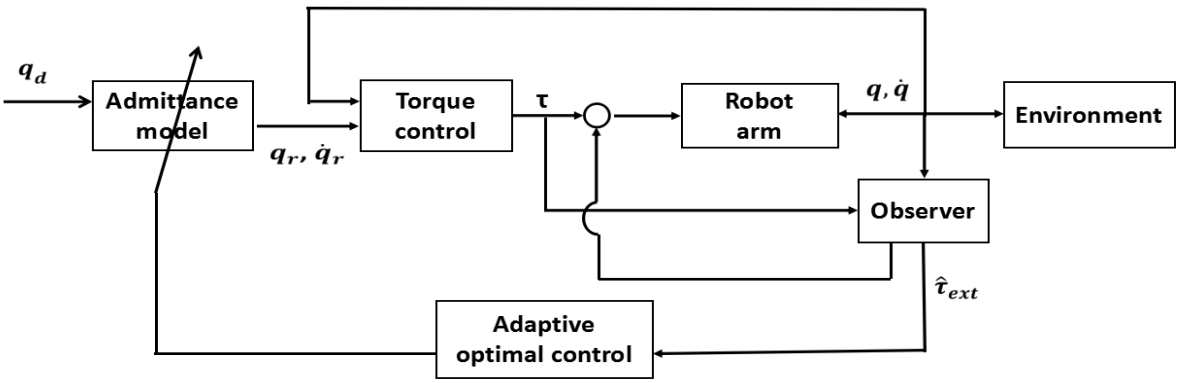

Fig. 2. The control diagram

environments can be represented by these two models. For analysis convenience, the spring-damper system is considered as a time-invariant system i.e. the three coefficient matrices are constant matrices.

B. Control Strategy

The system diagram is given in Fig. 2. The outer-loop of the system is to obtain admittance parameters subject to an unknown environment based on the adaptive optimal control method. With the interaction torque estimated from an observer, the virtual desired trajectory q_r in the joint space is generated. The inner-loop of the system is to guarantee the trajectory tracking with adaptive control scheme.

In general, the desired admittance model in the Cartesian space is

$$f_{ext} = f(x_r, x_d) \tag{6}$$

where $x_r \in \mathbb{R}^n$ is the virtual desired trajectory in the Cartesian space and $f(\cdot)$ is the function of the admittance model. To be specific, a target admittance model in joint space is described as below

$$\begin{aligned} M_d J(q)(\ddot{q}_r - \ddot{q}_d) + (M_d \dot{J}(q) + C_d J(q))(\dot{q}_r - \dot{q}_d) \\ + G_d(\kappa(q_r) - \kappa(q_d)) = -J(q)^{\mathrm{T}} \tau_{ext} \end{aligned} \tag{7}$$

where M_d, C_d and G_d are the desired inertia, damping and stiffness matrices, respectively.

Remark 2: Model (7) is a general admittance model which defines the relationship between interaction torque and joint angles. In certain situations, a more simplified stiffness model may be adopted

$$G_d(q_r - q_d) = -\tau_{ext} \tag{8}$$

Besides, M_d, C_d and G_d are constant matrices which implies identical characteristics in all directions. Obviously, in order to have different characteristics in different directions, these three matrices should be defined as positive definite matrices with different diagonal elements. When $\tau_{ext} = 0$, q_d is the desired trajectory to be tracked by in the absence of interaction. However, when τ_{ext} is not null, the virtual desired trajectory q_r will be generated.

The adaptive control scheme is to let robot follow the desired trajectory and drive the tracking error $e_q = q - q_r$ into a small neighborhood of zero. The design of adaptive control scheme will be discussed in the next section.

C. Control Objective

The control objective is to achieve an optimal interaction performance and the following cost function is defined to quantify the interaction performance

$$V(t) = \int_0^\infty \left((q - q_d)^{\mathrm{T}} Q (q - q_d) + \hat{\tau}_{ext}^{\mathrm{T}} R \hat{\tau}_{ext} \right) dt \tag{9}$$

where $Q = Q^{\mathrm{T}} \in \mathbb{R}^{n \times n}$ is positive definite, describing the weight of tracking errors, and $R \in \mathbb{R}^{n \times n}$ is the weight of the interaction torque. By minimizing $V(t)$, a desired interaction performance can be achieved.

Remark 3: Different cost functions similar to (9) have been discussed in some related works [21]. In a traditional LQR problem, a cost function includes control input and trajectory tracking errors. The optimal control performance can be achieved by specifying feedback gains. In this paper, both the robot and environment systems are taken into consideration.

III. METHODOLOGY

The objective of the admittance adaptation is to obtain the target admittance model according the changing environment. The inner-loop of the system is to guarantee the tracking performance. The interaction torque from the environment is estimated by the force observer.

A. Torque Estimation

In this section, an observer based on the generalized momentum approach is used to estimate the external torque in joint space. Compared with traditional method requiring computation of joint accelerations or the inversion of the inertia matrix [33], the generalized momentum approach assumes that only motor torque τ, joint angle q and joint velocity $\dot{q}$ are available. In [33] the generalized momentum is defined as

$$p = M(q)\dot{q} \tag{10}$$

Its differential form with respect to time is

$$\dot{p} = \dot{M}\dot{q} + M\ddot{q} \tag{11}$$

Substituting (11) into (4), we have

$$\dot{p} = \dot{M}(q, \dot{q})\dot{q} + \tau - C(q, \dot{q})\dot{q} - G(q) - \tau_{ext} \tag{12}$$

Considering that the matrix M is symmetric and positive definite and Coriolis matrix is expressed using Christoffel

symbols [32], the time derivative of inertia matrix M can be written as

$$\dot{M} = C + C^{\mathrm{T}} \tag{13}$$

Substituting (13) into (12), we have

$$\dot{p} = C^{\mathrm{T}}(q, \dot{q})\dot{q} + \tau - G(q) - \tau_{ext} \tag{14}$$

It is obvious that equation (14) based on the generalized momentum does not involve joint angle acceleration $\ddot{q}$. Finally, the external torques can be modelled as

$$\dot{\tau}_{ext} = A_\tau \tau_{ext} + w_\tau \tag{15}$$

where w_τ is the uncertainty, $w_\tau \sim N(0, Q_\tau)$. Usually, the matrix A_τ is defined as $A_\tau = 0_{n\times n}$. However, a negative diagonal matrix can reduce the offset of the estimation of disturbances. Then, equation (14) can be rewritten as

$$\dot{p} = u - \tau_{ext} \tag{16}$$

where u is defined as

$$u = \tau + C^{\mathrm{T}}(q, \dot{q})\dot{q} - G(q) \tag{17}$$

The above equations can be combined and reformulated in the state-space form

$$\begin{aligned} \underbrace{\begin{bmatrix} \dot{p} \\ \dot{\tau}_{ext} \end{bmatrix}}_{\dot{x}} &= \underbrace{\begin{bmatrix} 0_n & -I_n \\ 0_n & A_\tau \end{bmatrix}}_{A_c} \underbrace{\begin{bmatrix} p \\ \tau_{ext} \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} I_n \\ 0_n \end{bmatrix}}_{B_c} u + \underbrace{\begin{bmatrix} 0 \\ w_\tau \end{bmatrix}}_{w} \\ y &= \underbrace{\begin{bmatrix} I_n & 0_n \end{bmatrix}}_{C_c} \begin{bmatrix} p \\ \tau_{ext} \end{bmatrix} + v \end{aligned} \tag{18}$$

where v is the measurement noise $v \sim N(0, R_c)$. It can be easily proved that this system is observable. Since q and $\dot{q}$ are able to be measured, the generalized momentum $p = M(q)\dot{q}$ can be regarded as a measured variable. Then, a state observer is designed

$$\begin{cases} \dot{\hat{x}} = A_c\hat{x} + B_c u + L(y - \hat{y}) \\ \hat{y} = C_c\hat{x} \end{cases} \tag{19}$$

The gain matrix L in the system can be calculated by

$$L = P{C_c}^{\mathrm{T}} R_c^{-1} \tag{20}$$

where the matrix P can be calculated by the algebraic Riccati equation (ARE) [34]

$$A_c P + P{A_c}^{\mathrm{T}} - P{C_c}^{\mathrm{T}} R_c^{-1} C_c P + Q_c = 0 \tag{21}$$

where Q_c is the uncertainty of the state, written as

$$Q_c = diag([0, Q_\tau]) \tag{22}$$

A schematic overview of the observer is shown in Fig. 3. As shown in equation (19), the output $y = C_c x(t)$ is compared with $C_c\hat{x}(t)$. If the gain matirx L is properly designed, the difference, passing through the gain matrix, will drive the estimated state to actual state. From the above analysis, we can see that the estimation of states can be obtained from estimated state $\hat{x}$, which can be written as

$$\begin{aligned} \hat{\tau}_{ext} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \hat{x} \\ \hat{p} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \hat{x} \end{aligned} \tag{23}$$

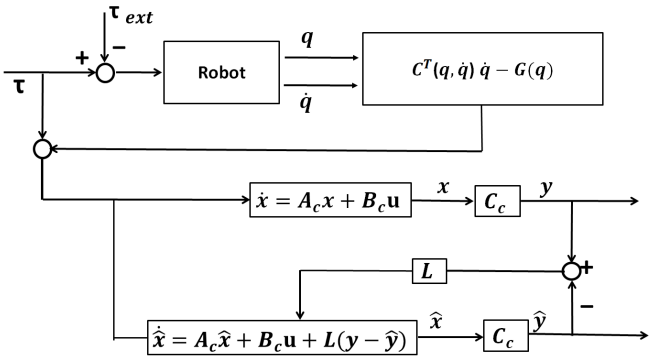

Fig. 3. The diagram of the observer based on generalized momentum approach

B. Adaptive Optimal Control

The adaptive optimal control strategy is proposed in [31], which is outlined in the following. Consider a continuous-time linear system:

$$\dot{\xi}(t) = A\xi(t) + Bu(t) \tag{24}$$

where $\xi \in \mathbb{R}^m$ is the system state variable, $u \in \mathbb{R}^r$ is the system input, $A \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{m\times r}$ are the system matrix and input matrix assumed to be constant and unknown. The following optimal control input

$$u = -K\xi \tag{25}$$

which can minimize the cost function as follows

$$V = \int_0^\infty \left(\xi^{\mathrm{T}} Q \xi + u^{\mathrm{T}} R\, u\right) dt \tag{26}$$

The solution to this problem is similar to that of the LQR problem. The LQR provides a systematic way to find feedback gains that guarantee the optimal control performance. In optimal control theory [35], when A and B are known, there exists a symmetric positive matrix P^*, which is the solution of the ARE

$$PA + A^{\mathrm{T}}P + Q - PBR^{-1}B^{\mathrm{T}}P = 0 \tag{27}$$

Then, we can obtain the optimal feedback gain matrix

$$K^* = -R^{-1}B^{\mathrm{T}}P^* \tag{28}$$

Therefore, we can obtain the optimal control input from equation (28). Next, we give the on-line learning algorithm to obtain the optimal control input subject to unknown dynamics of the environment.

For convenience, let us introduce the following definitions: $P_k \in \mathbb{R}^{m\times m} \to \hat{P}_k \in \mathbb{R}^{\frac{1}{2}m(m+1)}$ and $\xi \in \mathbb{R}^m \to \mathbb{R}^{\frac{1}{2}m(m+1)}$ where P_k is a symmetric matrix

$$\begin{aligned} \hat{P}_k &= [P_{11},\ 2P_{12}, ...,\ 2P_{1m},\ P_{22},\ 2P_{23}, ...,\ P_{mm}]^{\mathrm{T}} \\ \bar{\xi} &= [\xi_1^2,\ \xi_1\xi_2, ...,\ \xi_1\xi_m,\ \xi_2^2,\ \xi_2\xi_3, ...,\ \xi_m^2]^{\mathrm{T}} \\ \delta_{\xi\xi} &= [\bar{\xi}(t_1) - \bar{\xi}(t_0),\ \bar{\xi}(t_2) - \bar{\xi}(t_1),\ ....,\bar{\xi}(t_l) - \bar{\xi}(t_{l-1})]^{\mathrm{T}} \\ I_{\xi\xi} &= \left[\int_{t_0}^{t_1} \xi\otimes\xi dt,\ \int_{t_1}^{t_2} \xi\otimes\xi dt, ...,\ \int_{t_{l-1}}^{t_l} \xi\otimes\xi dt\right]^{\mathrm{T}} \\ I_{\xi u} &= \left[\int_{t_0}^{t_1} \xi\otimes u dt,\ \int_{t_1}^{t_2} \xi\otimes u dt, ...,\ \int_{t_{l-1}}^{t_l} \xi\otimes u dt\right]^{\mathrm{T}} \end{aligned} \tag{29}$$

where l is a positive integer and $\otimes$ is the Kronecker product. Now, we let $u = K_0\xi + \phi$ be the initial input, $t \in [t_0,\ t_l]$. ϕ is the exploration noise and K_0 is initial feedback gain, which can stabilize the system. Then, compute $I_{\xi\xi}$ and $I_{\xi u}$ until the following rank condition is satisfied

$$rank\left([I_{\xi\xi},\ I_{\xi u}]\right) = \frac{m(m+1)}{2} + mr \tag{30}$$

After the rank condition is satisfied, we can solve P_k and K_{k+1} according to the following equation

$$\begin{bmatrix} \hat{P}_k \\ vec(K_{k+1}) \end{bmatrix} = (\Theta_k^{\mathrm{T}}\Theta_k)^{-1}\Theta_k^{\mathrm{T}}\Xi_k \tag{31}$$

Θ_k and Ξ_k are defined as

$$\begin{aligned} \Theta_k &= [\delta_{\xi\xi}, -2I_{\xi\xi}(I_m \otimes K_k^{\mathrm{T}}R) - 2I_{\xi u}(I_m \otimes R)] \\ \Xi_k &= -I_{\xi\xi}vec(Q_k) \\ Q_k &= Q + K_k^{\mathrm{T}}RK_k \end{aligned} \tag{32}$$

where I_m is the m-dimensional unit matrix, $vec(\cdot)$ is the function to transfer a matrix to a vector. Then, we repeat the calculation until $||P_k - P_{k-1}|| < \varepsilon$, where ε is a small constant defined by the designer. Finally, the optimal feedback gain K_k is obtained. This learning algorithm is summarized in Algorithm 1.

C. Admittance Adaptation

This section is to obtain a target admittance model according to an unknown environment. The environment model is assumed to be damping-stiffness as described in equation (5). The cost function defined in the equation (9) is to be minimised. Comparing the cost function in this paper with the general linear system (24), we need to make them identical. Define the state variable

$$\xi = [q^{\mathrm{T}}, q_d^{\mathrm{T}}]^{\mathrm{T}} \tag{33}$$

Then, equation (9) can be

$$\begin{aligned} V &= \int_0^\infty \left([q^{\mathrm{T}} q_d^{\mathrm{T}}] Q' \begin{bmatrix} q \\ q_d \end{bmatrix} + \hat{\tau}_{ext}^{\mathrm{T}}\ R\ \hat{\tau}_{ext} \right) dt \\ &= \int_0^\infty \left(\xi^{\mathrm{T}} Q' \xi + \hat{\tau}_{ext}^{\mathrm{T}}\ R\ \hat{\tau}_{ext} \right) dt \end{aligned} \tag{34}$$

where

$$Q' = \begin{bmatrix} Q & -Q \\ -Q & Q \end{bmatrix} \tag{35}$$

Combined with the defined state variable, we can rewrite the environment model into state-space form

$$\dot{\xi} = A\xi + B\hat{\tau}_{ext} \tag{36}$$

where

$$A = \begin{bmatrix} -C_E^{-1}G_E & 0 \\ 0 & I_n \end{bmatrix}, B = \begin{bmatrix} -C_E^{-1} \\ 0 \end{bmatrix} \tag{37}$$

It is obvious that matrices A and B contain the unknown dynamics of the environment. If $\hat{\tau}_{ext}$ is taken as the input of the system (36), we can use the adaptive optimal control method discussed above to obtain the control input as follows to minimise the cost function

$$\hat{\tau}_{ext} = -K_k\xi \tag{38}$$

Algorithm 1 Admittance Adaptation Algorithm

1: Choose $u = K_0\xi + \phi$ as the initial admittance model, where K_0 is the initial feedback gain and ϕ is the exploration noise. Compute $\delta_{\xi\xi}, I_{\xi\xi}, I_{\xi u}$ until the rank condition in equation (30) is satisfied.
2: Solve P_k and K_{k+1} in equation (32).
3: Let $K+1 \rightarrow K$ and repeat Step 2 until $||P_k - P_{k-1}|| < \varepsilon$, where ε is a small constant.
4: Use $u = K_k\xi$ as the approximated optimal control input.

where K_k will be obtained by the on-line learning algorithm discussed in the previous section.

To understand equation (38) in the sense of LQR, the optimal control input is obtained in equation (28). According to the solution of ARE, we can obtain the optimal matrix

$$P^* = \begin{bmatrix} P_1 & P_2 \\ * & * \end{bmatrix} \tag{39}$$

where $P_1 \in \mathbb{R}^{n\times n}$ and $P_2 \in \mathbb{R}^{n\times n}$ and $*$ denotes the useless matrix. Therefore, we have

$$\hat{\tau}_{ext} = R^{-1}P_1 q - R^{-1}P_2 q_d \tag{40}$$

Comparing equation (40) with the desired admittance model (6), the expected admittance model is obtained to ensure the optimal interaction performance. With the desired trajectory q_d and $\hat{\tau}_{ext}$ estimated by the observer approach, we can obtain the virtual desired trajectory q_r in joint space and the inner-loop is to guarantee the trajectory tracking.

D. RBFNN

Radial basis function neural network (RBFNN) has capabilities of approximating any continous function [36] $f(\theta)$: $R^m \rightarrow R$ as follows

$$f(\theta) = W^{\mathrm{T}}Z(\theta) + \varepsilon(\theta) \tag{41}$$

where the input vector $\theta \in \Omega_\theta \subset R^m$; weight vector $W = [w_1, w_2, ..., w_d] \in R^d$, d denotes the NNs node number $d > 1$; $Z(\theta) = [Z_1(\theta), Z_2(\theta), ..., Z_T(\theta)]^{\mathrm{T}}$, with $Z_i(\theta)$ the basis function usually chosen as Gaussian function as

$$Z_i(\theta) = exp[\frac{-(\theta - u_i^{\mathrm{T}})(\theta - u_i)}{\eta_i^2}],\ i = 1, ..., d \tag{42}$$

where $u_i = [u_{i1}, u_{i2}, ..., u_{im}]^{\mathrm{T}} \in R^m$ is the center field and η_i the standard deviation. A continuous function can be approximated by

$$f(\theta) = W^{*\mathrm{T}}Z(\theta) + \varepsilon^*(\theta) \tag{43}$$

The ideal weight vector W^* is defined as the value of W which minimizes $\varepsilon(\theta)$ for all $\theta \in \Omega_\theta \subset R^m$

$$W^* = arg\min_{W\subset R^t}\{sup|f(\theta) - W^{\mathrm{T}}Z(\theta)|\} \tag{44}$$

In general, the ideal weights W^* are unknown and need to be estimated by $\hat{W}$. The weight estimation errors are defined as

$$\tilde{W} = W^* - \hat{W} \tag{45}$$

E. Controller Design

The inner-loop of the system is to guarantee tracking performance. An adaptive neural based controller is designed to achieve the objective. Considering the dynamics of robot manipulator (4), we define

$$\begin{aligned} s &= \dot{e}_q - \Lambda e_q \\ v &= \dot{q}_r + \Lambda e_q \\ \Lambda &= diag(\lambda_1, \lambda_2, .., \lambda_n) \end{aligned} \tag{46}$$

where $e_q = q - q_r$ and $\lambda_i > 0$. Substituting (46) into (4), we have

$$M(q)\dot{s}+C(q,\dot{q})s+G(q)+M(q)\dot{v}+C(q,\dot{q})v = \tau-\tau_{ext} \tag{47}$$

Design the control torque input

$$\tau = \hat{G} + \hat{M}\dot{v} + \hat{C}v + \hat{\tau}_{ext} - Ks \tag{48}$$

where $\hat{G}(q)$, $\hat{M}(q)$ and $\hat{C}(q,\dot{q})$ are the estimates of $G(q)$, $M(q)$ and $C(q,\dot{q})$; $K = diag(k_1, k_2, ..., k_i)$ with $k_i > \frac{1}{2}$. The closed-loop dynamics can be written as

$$\begin{aligned} M(q)\dot{s} + C(q,\dot{q})s + Ks = &- (M - \hat{M})\dot{v} - (C - \hat{C})v \\ &- (G - \hat{G}) + (\hat{\tau}_{ext} - \tau_{ext}) \end{aligned} \tag{49}$$

Using the approximation method, we have

$$\begin{aligned} M(q) &= W_M^{\mathrm{T}} Z_M(q) + \varepsilon_M(q) \\ C(q,\dot{q}) &= W_C^{\mathrm{T}} Z_C(q,\dot{q}) + \varepsilon_C(q) \\ G(q) &= W_G^{\mathrm{T}} Z_G(q) + \varepsilon_G(q) \end{aligned} \tag{50}$$

where W_M, W_C, W_G are the ideal weight matrices; $Z_M(q)$, $Z_C(q,\dot{q})$, $Z_G(q)$ are the basis function matrices. The basis function matrices can be defined as

$$\begin{aligned} Z_M(q) &= diag(Z_q, ..., Z_q) \\ Z_C(q,\dot{q}) &= diag([Z_q, Z_{\dot{q}}]^{\mathrm{T}}, ..., [Z_q, Z_{\dot{q}}]^{\mathrm{T}}) \\ Z_G(q) &= diag(Z_q^{\mathrm{T}}, ..., Z_q^{\mathrm{T}}) \end{aligned} \tag{51}$$

where

$$\begin{aligned} Z_q &= [\psi(||q - q_1||), ..., \psi(||q - q_n||)]^{\mathrm{T}} \\ Z_{\dot{q}} &= [\psi(||\dot{q} - \dot{q}_1||), ..., \psi(||\dot{q} - \dot{q}_n||)]^{\mathrm{T}} \end{aligned} \tag{52}$$

and $\psi(\cdot)$ is defined as the Gaussian function. The estimates of $M(q)$, $C(q,\dot{q})$ and $G(q)$ can be written as

$$\begin{aligned} \hat{M}(q) &= \hat{W}_M^{\mathrm{T}} Z_M(q) \\ \hat{C}(q,\dot{q}) &= \hat{W}_C^{\mathrm{T}} Z_C(q,\dot{q}) \\ \hat{G}(q) &= \hat{W}_G^{\mathrm{T}} Z_G(q) \end{aligned} \tag{53}$$

Substituting (50) and (53) into (49), we have

$$\begin{aligned} M(q)\dot{s} + C(q,\dot{q})s + Ks = &- \tilde{W}_M^{\mathrm{T}} Z_M \dot{v} - \tilde{W}_C^{\mathrm{T}} Z_C v \\ &- \tilde{W}_G^{\mathrm{T}} Z_G - e_\tau \end{aligned} \tag{54}$$

where $\tilde{W}_M = W_M^* - \hat{W}_M, \tilde{W}_C = W_C^* - \hat{W}_C$ and $\tilde{W}_G = W_G^* - \hat{W}_G$ and $e_\tau = \tau_{ext} - \hat{\tau}_{ext}$. Considering the Lyapunov function

$$V = \frac{1}{2}s^{\mathrm{T}}Ms + \frac{1}{2}tr(\tilde{W}_M^{\mathrm{T}} Q_M \tilde{W}_M + \tilde{W}_C^{\mathrm{T}} Q_C \tilde{W}_C + \tilde{W}_G^{\mathrm{T}} Q_G \tilde{W}_G) \tag{55}$$

where Q_M, Q_C,Q_G are positive definite matrices to be set by the designer. The derivative of V can be written as

$$\begin{aligned} \dot{V} =& s^{\mathrm{T}} M\dot{s} + \frac{1}{2}s^{\mathrm{T}}\dot{M}s \\ &+ tr(\tilde{W}_M^{\mathrm{T}} Q_M \dot{\tilde{W}}_M + \tilde{W}_C^{\mathrm{T}} Q_C \dot{\tilde{W}}_C + \tilde{W}_G^{\mathrm{T}} Q_G \dot{\tilde{W}}_G) \end{aligned} \tag{56}$$

By using the estimation of the weight matrices $\hat{W}_M, \hat{W}_C, \hat{W}_G$ to approximate W_M^*, W_C^*, W_G^*, the errors between the actual and the ideal RBFNN can be expressed as

$$\begin{aligned} W_M^{*\mathrm{T}} Z_M(q) - \hat{W}_M^{\mathrm{T}} Z_M(q) &= \tilde{W}_M^{\mathrm{T}} Z_M(q) \\ W_C^{*\mathrm{T}} Z_C(q) - \hat{W}_C^{\mathrm{T}} Z_C(q) &= \tilde{W}_C^{\mathrm{T}} Z_C(q) \\ W_G^{*\mathrm{T}} Z_G(q) - \hat{W}_G^{\mathrm{T}} Z_G(q) &= \tilde{W}_G^{\mathrm{T}} Z_G(q) \end{aligned} \tag{57}$$

As the ideal weight matrix W^* is a constant vector, we know that

$$\begin{aligned} \dot{\tilde{W}}_M &= \dot{\hat{W}}_M \\ \dot{\tilde{W}}_C &= \dot{\hat{W}}_C \\ \dot{\tilde{W}}_G &= \dot{\hat{W}}_G \end{aligned} \tag{58}$$

Considering $2C(q,\dot{q}) - \dot{M}(q)$ is a skew-symmetric matrix [32] and (54), we have

$$\begin{aligned} \dot{V} =& s^{\mathrm{T}} M\dot{s} + s^{\mathrm{T}} C\dot{s} \\ &+ tr(\tilde{W}_M^{\mathrm{T}} Q_M \dot{\tilde{W}}_M + \tilde{W}_C^{\mathrm{T}} Q_C \dot{\tilde{W}}_C + \tilde{W}_G^{\mathrm{T}} Q_G \dot{\tilde{W}}_G) \\ =& - s^{\mathrm{T}}(Ks + \tilde{W}_M^{\mathrm{T}} Z_M \dot{v} + \tilde{W}_C^{\mathrm{T}} Z_C v + \tilde{W}_G^{\mathrm{T}} Z_G + e_\tau) \\ &+ tr(\tilde{W}_M^{\mathrm{T}} Q_M \dot{\tilde{W}}_M + \tilde{W}_C^{\mathrm{T}} Q_C \dot{\tilde{W}}_C + \tilde{W}_G^{\mathrm{T}} Q_G \dot{\tilde{W}}_G) \\ =& - s^{\mathrm{T}} Ks - s^{\mathrm{T}} e_\tau \\ &- tr\left[\tilde{W}_M^{\mathrm{T}}(Z_M \dot{v} s^{\mathrm{T}} + Q_M \dot{\tilde{W}}_M)\right] \\ &- tr\left[\tilde{W}_C^{\mathrm{T}}(Z_C \dot{v} s^{\mathrm{T}} + Q_C \dot{\tilde{W}}_C)\right] \\ &- tr\left[\tilde{W}_C^{\mathrm{T}}(Z_G s^{\mathrm{T}} + Q_G \dot{\tilde{W}}_G)\right] \end{aligned} \tag{59}$$

The update law is designed as

$$\begin{aligned} \dot{\hat{W}}_M &= -Q_M^{-1}(Z_M \dot{v} s^{\mathrm{T}} + \sigma_M \hat{W}_M) \\ \dot{\hat{W}}_C &= -Q_C^{-1}(Z_C \dot{v} s^{\mathrm{T}} + \sigma_C \hat{W}_C) \\ \dot{\hat{W}}_G &= -Q_G^{-1}(Z_G \dot{v} s^{\mathrm{T}} + \sigma_G \hat{W}_G) \end{aligned} \tag{60}$$

where σ_M, σ_G and σ_C are constants to be specified by the designer. Substituting (60) into (59), the derivative of V is

$$\begin{aligned} \dot{V} =& - s^{\mathrm{T}} Ks - s^{\mathrm{T}} e_\tau + tr[\sigma_M \tilde{W}_M^{\mathrm{T}} \hat{W}_M] \\ &+ tr[\sigma_C \tilde{W}_C^{\mathrm{T}} \hat{W}_C] + tr[\sigma_G \tilde{G}_M^{\mathrm{T}} \hat{W}_G] \end{aligned} \tag{61}$$

Using Young's inequality [37]

$$\begin{aligned} tr[\tilde{W}_{(\cdot)}^{\mathrm{T}} \hat{W}_{(\cdot)}] &\leq -\frac{1}{2}||\tilde{W}_{(\cdot)}||_F^2 + \frac{1}{2}||W_{(\cdot)}^*||_F^2 \\ -s^{\mathrm{T}} e_\tau &\leq \frac{1}{2}s^{\mathrm{T}} s + \frac{1}{2}e_\tau^{\mathrm{T}} e_\tau \end{aligned} \tag{62}$$

Then (61) can be written as

$$\begin{aligned} \dot{V} \leq &- s^{\mathrm{T}} Ks + \frac{1}{2}||s||^2 + \frac{1}{2}||e_\tau||^2 + \alpha \\ &- \frac{\sigma_M}{2}||\tilde{W}_M||^2 - \frac{\sigma_C}{2}||\tilde{W}_C||^2 - \frac{\sigma_G}{2}||\tilde{W}_G||^2 \end{aligned} \tag{63}$$

where $\alpha = \frac{\sigma_M}{2}||W_M^*||^2 + \frac{\sigma_C}{2}||W_C^*||^2 + \frac{\sigma_G}{2}||W_G^*||^2$. A sufficient condition for $\dot{V} \leq 0$ is that

$$\begin{aligned}\alpha \leq & s^{\mathrm{T}}(K - \frac{1}{2}I)s + \frac{1}{2}||e_\tau||^2 + \frac{\sigma_M}{2}||\tilde{W}_M||^2 \\ & + \frac{\sigma_C}{2}||\tilde{W}_C||^2 + \frac{\sigma_G}{2}||\tilde{W}_G||^2\end{aligned} \tag{64}$$

where I is the unit matrix. Let χ denotes the state variable comprised of $e_\tau, s, \tilde{W}_M, \tilde{W}_C, \tilde{W}_G$ defined in the Lyapunov function candidate, and it follows from (64) that

$$\dot{V}(\chi) < 0, \forall ||\chi|| > \varrho \tag{65}$$

where ϱ is a positive constant. In other words, the time derivative of $V(\chi)$ is negative outside the set $\Omega_s = \{||\chi|| \leq \varrho\}$, which is defined in Theorem 1 as below, or equivalently, all $\chi(t)$ that start outside Ω_s will enter the set within a finite time, and will remain inside the set afterwards. Choose $0 < V(\chi) < \epsilon < c$, and suppose that the sets $\Omega_s = \{V(\chi) \leq \epsilon\}$ and $\Omega_c = \{V(\chi) \leq c\}$. Let

$$\Upsilon = \{\epsilon \leq V(\chi) \leq c\} = \Omega_c - \Omega_s \tag{66}$$

It is known that the time derivative of $V(\chi)$ is negative inside Υ, that is

$$\dot{V}(\chi(t)) < 0, \forall \chi \in \Upsilon, \forall t \geq t_0 \tag{67}$$

Since $\dot{V}$ is negative in $\Upsilon = \{\epsilon \leq V(\chi) \leq c\}$, which implies that in this set $V(\chi(t))$ will decrease monotonically in time until the solution enters the set $\{V(\chi) \leq \epsilon\}$. From that time on, $\chi(t)$ cannot leave the set because $\dot{V}$ is negative on its boundary $V(\chi) = \epsilon$. A sketch of the sets is shown in Fig. 4.

We can conclude the convergence and stability results in Theorem 1.

Theorem 1: Using the Uniformly Ultimately Bounded (UUB) theorem, the tracking error s, estimation error e_τ and weight errors $\tilde{W}_M, \tilde{W}_C, \tilde{W}_G$ will fall into an set Ω_s, where the bounding set Ω_s is defined in (68) and shown in Fig. 4.

$$\begin{aligned}\Omega_s = \Big\{ & (||\tilde{W}_M||, ||\tilde{W}_C||, ||\tilde{W}_G||, ||e_\tau||, ||s||), \mid \\ & \frac{\sigma_M||\tilde{W}_M||^2}{2\alpha} + \frac{\sigma_C||\tilde{W}_C||^2}{2\alpha} + \frac{\sigma_G||\tilde{W}_G||^2}{2\alpha} \\ & + \frac{s^{\mathrm{T}}(K - \frac{1}{2}I)s}{\alpha} + \frac{||e_\tau||^2}{\alpha} \leq 1 \Big\}\end{aligned} \tag{68}$$

As shown in Fig. 4, the bounding set Ω_s is the area in the first quadrant, passing through the points $(||\frac{\sigma_{(\cdot)}}{2}\tilde{W}_{(\cdot)}||^2 = \alpha, ||e_\tau||^2 = 0, ||s||^2 = 0)$, $(s^{\mathrm{T}}(K - \frac{1}{2}I)s = \alpha, ||\tilde{W}_{(\cdot)}||^2 = 0, ||e_\tau||^2 = 0)$ and $(\frac{1}{2}||e_\tau||^2 = \alpha, ||\tilde{W}_{(\cdot)}||^2 = 0, ||s||^2 = 0)$. In Fig. 4, we define

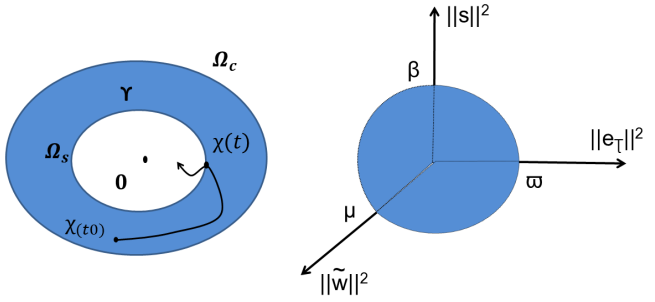

Fig. 4. Representation of UUB and the set Ω_s defined in (68).

$$\begin{aligned}& when \;\; ||\frac{\sigma_{(\cdot)}}{2}\tilde{W}_{(\cdot)}||^2 = \alpha, \;\; \tilde{W} = \varpi \\ & when \;\; s^{\mathrm{T}}(K - \frac{1}{2}I)s = \alpha, \;\; s = \beta \\ & when \;\; \frac{1}{2}||e_\tau||^2 = \alpha, \;\; e_\tau = \mu\end{aligned} \tag{69}$$

Since $||W_{(\cdot)}^*||$ is a bounded constant, $||\hat{W}_{(\cdot)}|| = ||W_{(\cdot)}^* - \tilde{W}_{(\cdot)}||$ is bounded. Since $||e_\tau||$ is bounded, $||\hat{\tau}_{ext}|| = ||e_\tau + \tau_{ext}||$ is bounded. With the bounded signals q_r and $\dot{q}_r$, according to (46), $||v||$ is bounded and $||q|| = ||e_q + q_r||$ is bounded as well. Therefore, the norm of state variable χ is bounded.

IV. Simulation Studies

In this section, we consider a 2-link robot arm in physical interaction with unknown environment. The simulation environment is shown in Fig. 5. The environment torque is applied at the end-effector of the robot arm in the Y direction. The desired trajectory will be modified to adjust the interaction torque exerted by the environment so that a compliant behavior is achieved. The controller is to guarantee the robot arm to track a virtual desired trajectory q_r. The parameters of the robot arm are shown in Table I.

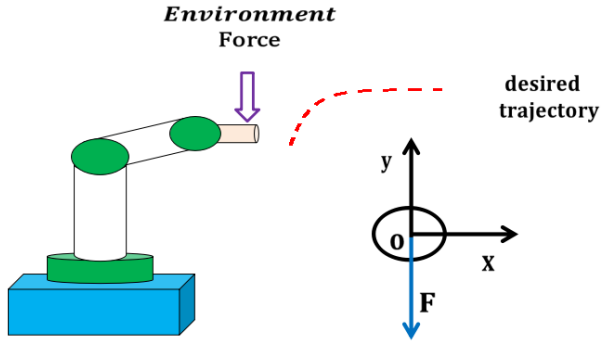

Fig. 5. An overview of the scenario.

TABLE I
Parameters of the Robot Arm

Parameters	Description	Value
m_1	Mass of link 1	2.0 kg
m_2	Mass of link 2	2.0 kg
l_1	Length of link 1	0.20 m
l_2	Length of link 2	0.20 m
I_1	Inertia moment of link 1	0.027 kgm^2
I_1	Inertia moment of link 2	0.027 kgm^2

A. Interaction Performance

It is assumed that the dynamics of the environment can be defined as: $0.01\dot{q} + (q - 0.3) = -\tau_{ext}$. The LQR method is used to verify the effectiveness of the proposed method. If the matrices A and B are known, we can use optimal solutions of the ARE to obtain the optimal parameters of the admittance

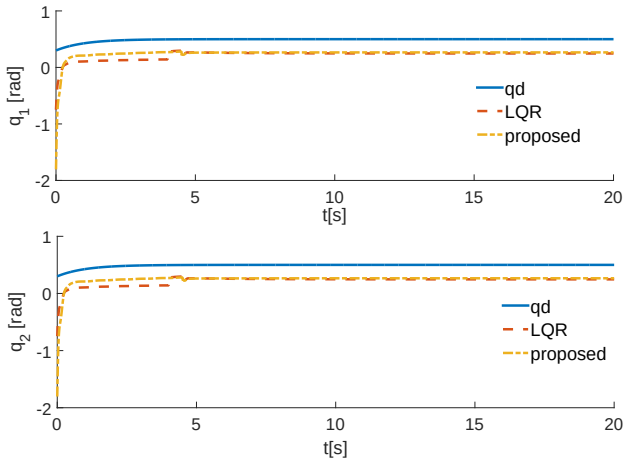

Fig. 6. The desired trajectory q_d and virtual desired trajectory q_r with $Q = 1$ and $R = 1$.

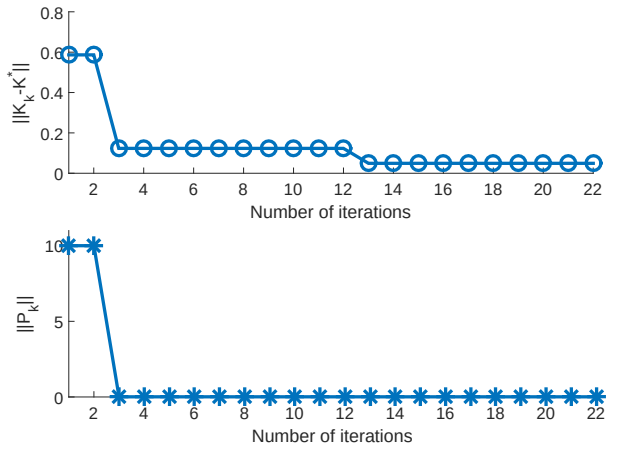

Fig. 7. The convergence of P_k and K_k to their optimal values with $Q = 1$ and $R = 1$.

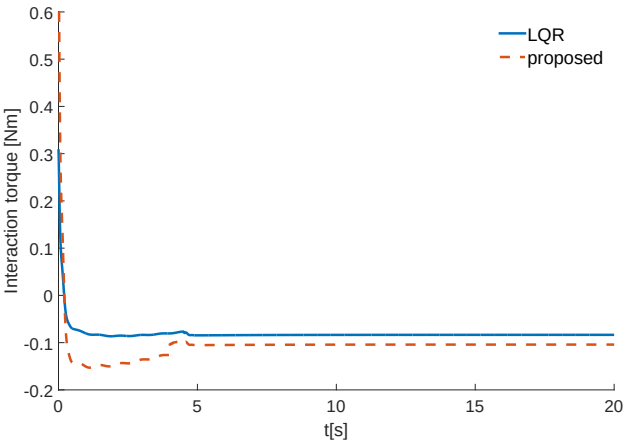

Fig. 8. Interaction torque with $Q = 1$ and $R = 1$.

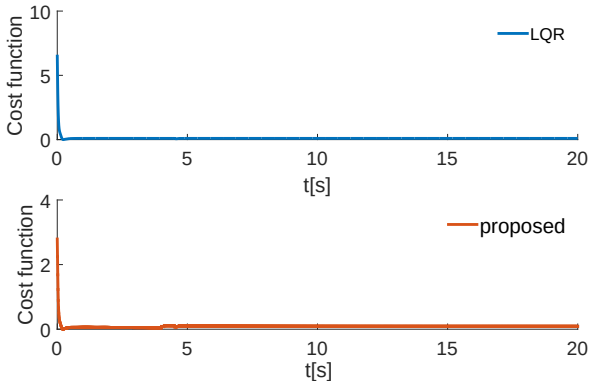

Fig. 9. The value of cost function with $Q = 1$ and $R = 1$.

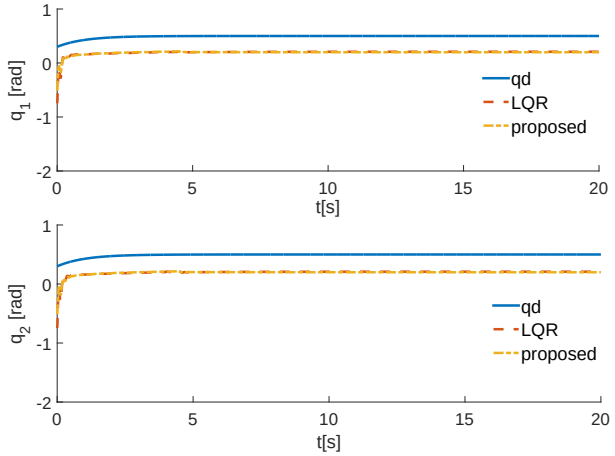

Fig. 10. The desired trajectory q_d and virtual desired trajectory q_r with $Q = 5$ and $R = 1$.

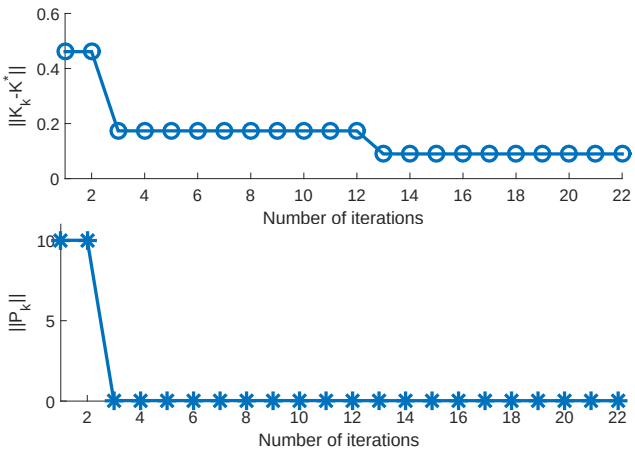

Fig. 11. The convergence of P_k and K_k to their optimal values with $Q = 5$ and $R = 1$.

model. The proposed method is adopted when the environment model is not available.

In the first step, the initial values of variables of the system should be properly set. The selection of exploration noise is not a trivial task which is used to make estimated parameters converge to the real values. In this paper, the exploration noise

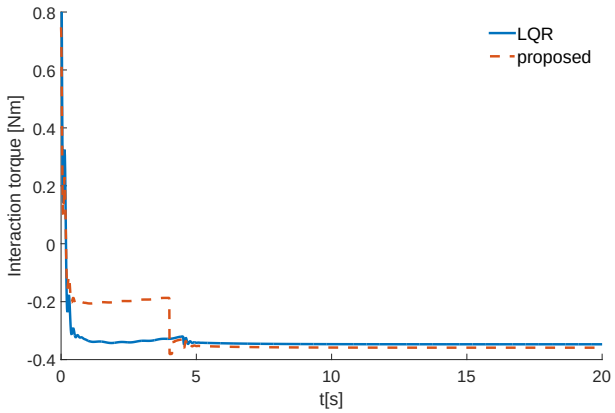

Fig. 12. Interaction torque with $Q = 5$ and $R = 1$.

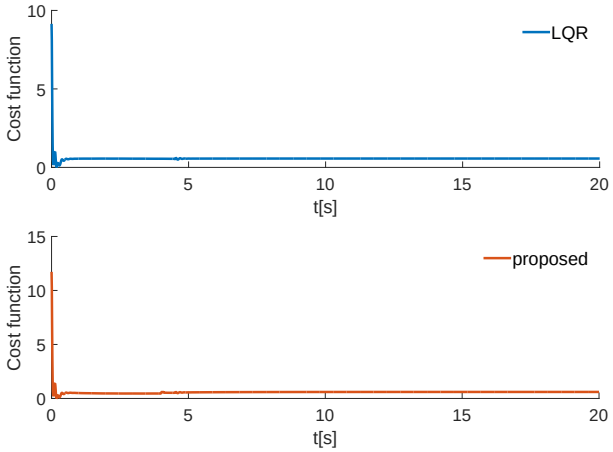

Fig. 13. The value of cost function with $Q = 5$ and $R = 1$.

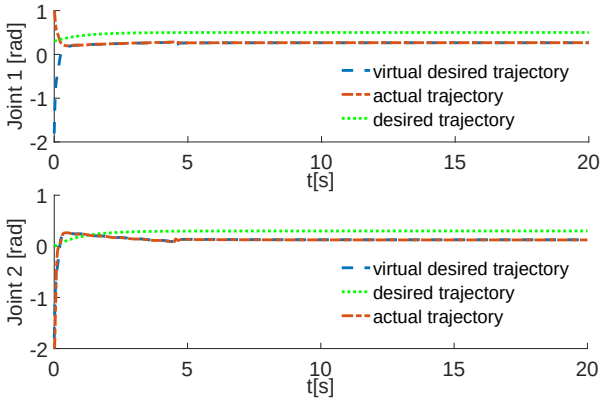

Fig. 14. The tracking performance of each joint.

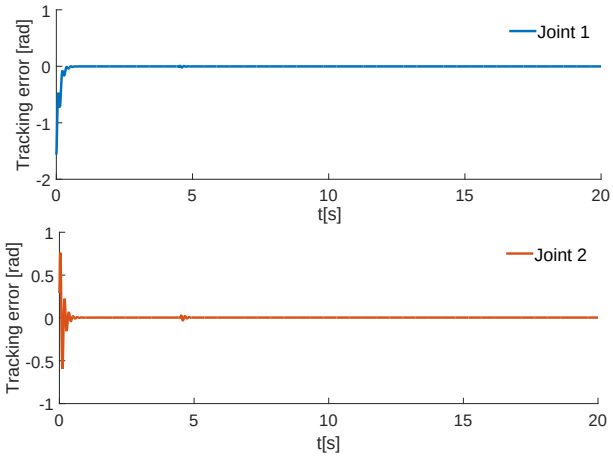

Fig. 15. Tracking error of each joint.

is selected as

$$\phi = \sum_{w=1}^{8} \frac{0.04}{w} sin(wt) \tag{70}$$

The initial feedback gain K_0 should ensure the stability of the system, which is set as $K_0 = [-1, 0.1]$. The initial P_k is set as $P_0 = 10I_p$, where I_p represents the p dimensional unit matrix. Then, the second step is conducted and stops until $||P_k|| < 0.02$, where $||\cdot||$ denotes the Euclidean norm.

The weights of the cost function are given by $Q = 1$ and $R = 1$. The desired admittance model is obtained as $\hat{\tau}_{ext} = -0.4142\dot{q}+0.0702q_r$ based on the known A and B. Simulation results are shown in Figs. 6-9. In Fig. 6, the desired trajectories of the robot arm in joint space are shown. At the beginning, there is a large error between the LQR and proposed method, due to the initial admittance model: $\hat{\tau}_{ext} = -\dot{q} + 0.1(q - 0.3)$ and the exploration noise. After that, the error is becoming smaller and trajectory of the robot arm with proposed method is coming closely to the trajectory of the robot arm with LQR method. The error of admittance parameters between LQR and proposed method is shown in Fig. 7. The convergence of P_k and K_k to optimal values based on LQR are illustrated. The error is defined as $||K_k - K^*||$. It is obvious that the error decreases to around 0.01 after 12 iterations and the Euclidean norm of P_k decreases to around 0.02. After three steps, the admittance model with the proposed method is obtained as: $\hat{\tau}_{ext} = -0.4173\dot{q} + 0.00904q_r$. The symmetric positive matrix with $Q = 1$ and $R = 1$, P_k from the proposed algorithm and P^* from LQR are shown below:

$$P_k = \begin{bmatrix} 0.0033 & 0.0027 \\ 0.0027 & 0.0077 \end{bmatrix} \quad P^* = \begin{bmatrix} 0.0041 & -0.0007 \\ -0.0007 & 0.0025 \end{bmatrix} \tag{71}$$

To further verify the correctness of the proposed method, different weights of the cost function are given by $Q = 5$ and $R = 1$. Simulation results are shown in Figs. 10-13. Similarly, the desired admittance model is obtained as $\hat{\tau}_{ext} = -1.4495\dot{q} + 0.2033q_r$ based on the known A and B. After three steps of iteration, the admittance model is obtained as $\hat{\tau}_{ext} = -1.5374\dot{q} + 0.2063q_r$ with the proposed method. The desired trajectory q_d and virtual desired trajectory q_r with $Q = 5$ and $R = 1$ are illustrated in Fig. 10. As shown in Fig. 11, the error $||K_k - K^*||$ between the LQR and proposed method converges to 0.08 after 12 iterations. The interaction torque converges to around $-0.35Nm$ at 7s, shown in Fig. 12. The values of cost function are 0.561 and 0.602 with LQR and the proposed method respectively. Similarly, the symmetric positive matrix with $Q = 5$ and $R = 1$, P_k from the proposed

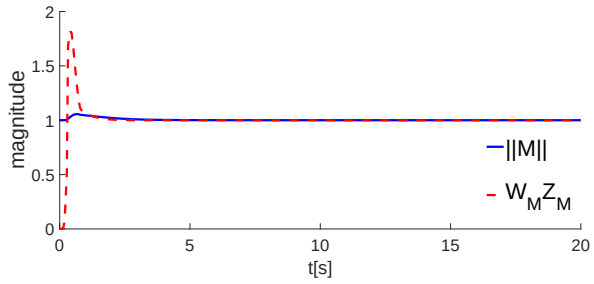

Fig. 16. Function approximation performance of the $M(q)$ matrix by RBF neural network.

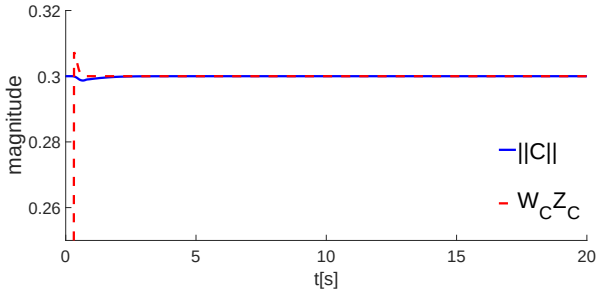

Fig. 17. Function approximation performance of the $C(q,\dot{q})$ matrix by RBF neural network.

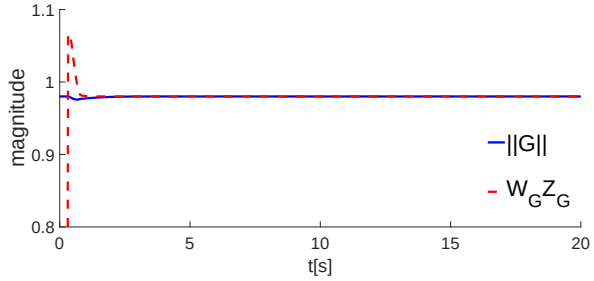

Fig. 18. Function approximation performance of the $G(q)$ matrix by RBF neural network.

algorithm and P^* from LQR are shown below:

$$P_k = \begin{bmatrix} 0.0106 & 0.0068 \\ 0.0068 & 0.0107 \end{bmatrix} \quad P^* = \begin{bmatrix} 0.0145 & -0.0020 \\ -0.0020 & 0.0043 \end{bmatrix} \tag{72}$$

Comparing the admittance model obtained by the proposed method and that with LQR method, the difference is small and acceptable. The overall results are satisfactory.

B. Trajectory Tracking

The initial joint angle of the robot arm is $q(0) = [2.6, -2.6]^{\mathrm{T}}$ and the desired trajectory of the robot arm is given by $q_d = [0.3 + 0.2e^{-t}, 0.3e^{-t}]^{\mathrm{T}}$. The weight matrices are initialized as $\hat{W}_M(0) = \mathbf{0}, \hat{W}_C(0) = \mathbf{0}, \hat{W}_G(0) = \mathbf{0}$. The results of tracking performance are shown in Figs. 14-20. In Fig. 14, the desired joint angle trajectory and actual trajectory are shown and the tracking errors are illustrated in Fig. 15. It can be found that the proposed control algorithm can make the tracking errors converge to a small neighborhood of zero. The

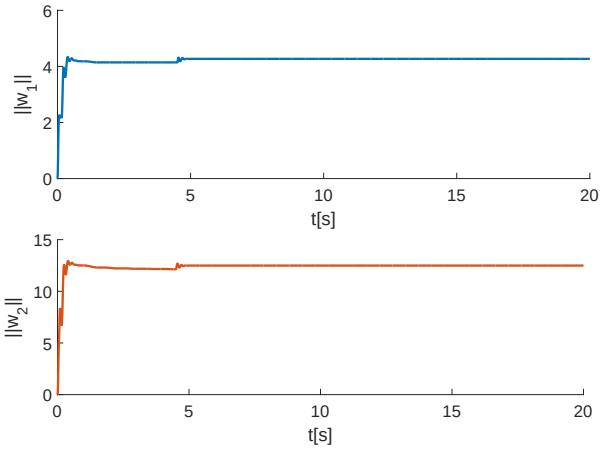

Fig. 19. The norm of weights of the neural network for each joint.

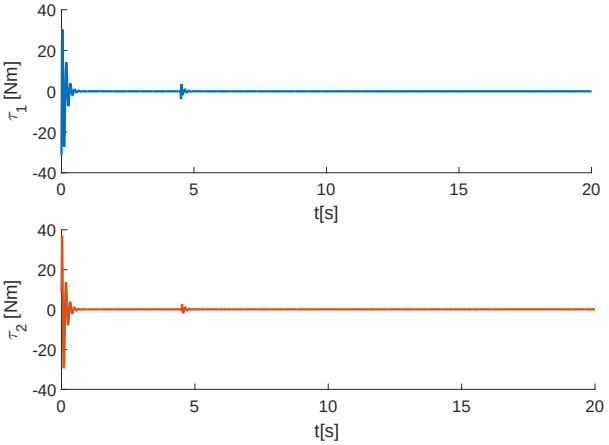

Fig. 20. The control inputs τ_1 and τ_2.

function approximation performance of the our implemented neural network is depicted in Figs. 16-18. As seen from the figures, we can observe that the output of the implemented neural network could follow the approximated nonlinear function's dynamics $(M(q), C(q,\dot{q}), G(q))$, which implies that the implemented neural networks have the ability to approximate the nonlinear function with satisfactory tracking performance. As shown in Fig. 19, the weight matrices of the system shows a trend of convergence and the control input is shown in Fig. 20. With the tracking errors and toque regulation, as shown in Fig. 8, the cost function is minimised, as shown in Fig. 9.

C. Torque Observer

The interaction torque is assumed to be applied at the end-effector of the robot arm. In Fig. 21, the actual interaction torque and its estimation are shown, and the estimated results of the generalized momentum of each joint are shown in Fig. 22. As seen from the figures, the output of the implemented torque observer could follow the real torque which implies that the implemented observer have a satisfactory estimation performance.

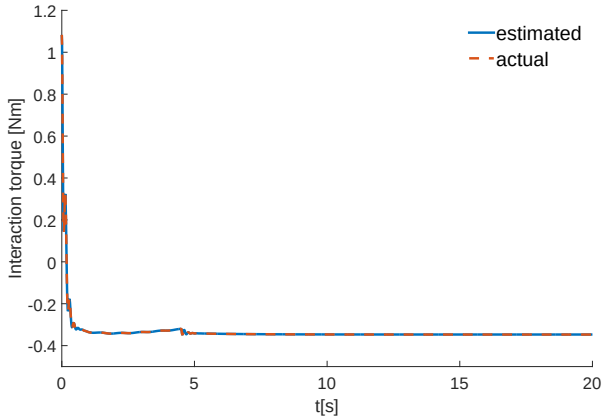

Fig. 21. The estimated interaction torque obtained from the observer based on generalized momentum approach and actual interaction torque.

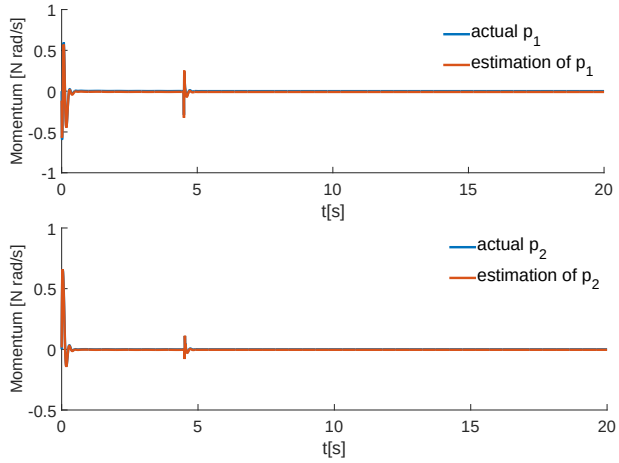

Fig. 22. The estimated generalized momentum of the joints obtained from the torque observer and actual generalized momentum.

V. CONCLUSION

In this paper, a method of admittance adaptation is proposed for robot-environment interaction. The desired admittance model is obtained by optimal adaptive control approach and admittance control is used to regulate the interaction behavior. A neural based controller is developed to guarantee trajectory tracking and interaction torque is estimated by an observer approach. A cost function that includes tracking errors and interaction torque is minimised. Simulation studies have verified the effectiveness of our proposed method.

REFERENCES

[1] S. Katsura and K. Ohnishi, "Human cooperative wheelchair for haptic interaction based on dual compliance control," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 1, pp. 221–228, 2004.

[2] H. Wang, B. Yang, Y. Liu, W. Chen, X. Liang, and R. Pfeifer, "Visual servoing of soft robot manipulator in constrained environments with an adaptive controller," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 41–50, Feb 2017.

[3] J. Huang, W. Huo, W. Xu, S. Mohammed, and Y. Amirat, "Control of upper-limb power-assist exoskeleton using a human-robot interface based on motion intention recognition," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1257–1270, Oct 2015.

[4] N. Hogan, "Impedance control: An approach to manipulation: Part iiłimplementation," *Journal of dynamic systems, measurement, and control*, vol. 107, no. 1, pp. 8–16, 1985.

[5] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 3, pp. 334–344, March 2016.

[6] M. T. Mason, "Compliance and force control for computer controlled manipulators," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 6, pp. 418–432, June 1981.

[7] A. M. Smith, C. Yang, H. Ma, P. Culverhouse, A. Cangelosi, and E. Burdet, "Novel hybrid adaptive controller for manipulation in complex perturbation environments," *PloS one*, vol. 10, no. 6, p. e0129281, 2015.

[8] H. Wang, D. Guo, H. Xu, W. Chen, T. Liu, and K. K. Leang, "Eye-in-hand tracking control of a free-floating space manipulator," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 4, pp. 1855–1865, Aug 2017.

[9] W. He and S. Zhang, "Control design for nonlinear flexible wings of a robotic aircraft," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 351–357, Jan 2017.

[10] W. He, Z. Yan, C. Sun, and Y. Chen, "Adaptive neural network control of a flapping wing micro aerial vehicle with disturbance observer," *IEEE Transactions on Cybernetics*, vol. 47, no. 10, pp. 3452–3465, Oct 2017.

[11] J. Huang, P. Di, T. Fukuda, and T. Matsuno, "Robust model-based online fault detection for mating process of electric connectors in robotic wiring harness assembly systems," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 5, pp. 1207–1215, Sept 2010.

[12] A. Alcocera, A. Robertssona, A. Valerac, and R. Johanssona, "Force estimation and control in robot manipulators," in *Robot Control 2003 (SYROCO'03): A Proceedings Volume from the 7th IFAC Symposium, Wrocław, Poland, 1-3 September 2003*, vol. 1. International Federation of Automatic Control, 2004, p. 55.

[13] A. De Luca and R. Mattone, "Sensorless robot collision detection and hybrid force/motion control," in *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005, pp. 999–1004.

[14] A. De Luca, A. Albu-Schaffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the dlr-iii lightweight manipulator arm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 1623–1630.

[15] M. Cohen and T. Flash, "Learning impedance parameters for robot control using an associative search network," *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 382–390, 1991.

[16] T. Tsuji, K. Ito, and P. G. Morasso, "Neural network learning of robot arm impedance in operational space," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 2, pp. 290–298, 1996.

[17] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li, and C. Y. Su, "Interface design of a physical human-robot interaction system for human impedance adaptive skill transfer," *IEEE Transactions on Automation Science and Engineering*, vol. PP, no. 99, pp. 1–12, 2017.

[18] R. Z. Stanisic and Á. V. Fernández, "Adjusting the parameters of the mechanical impedance for velocity, impact and force control," *Robotica*, vol. 30, no. 04, pp. 583–597, 2012.

[19] S. S. Ge, Y. Li, and C. Wang, "Impedance adaptation for optimal robot–environment interaction," *International Journal of Control*, vol. 87, no. 2, pp. 249–263, 2014.

[20] R. Johansson and M. W. Spong, "Quadratic optimization of impedance control," in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*. IEEE, 1994, pp. 616–621.

[21] M. Matinfar and K. Hashtrudi-Zaad, "Optimization-based robot compliance control: Geometric and linear quadratic approaches," *The International Journal of Robotics Research*, vol. 24, no. 8, pp. 645–656, 2005.

[22] D. E. Kirk, *Optimal control theory: an introduction*. Courier Corporation, 2012.

[23] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, MA, 1995, vol. 1, no. 2.

[24] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE circuits and systems magazine*, vol. 9, no. 3, pp. 32–50, 2009.

[25] F.-Y. Wang, H. Zhang, and D. Liu, "Adaptive dynamic programming: an introduction," *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 39–47, 2009.

[26] P. J. Werbos, "A menu of designs for reinforcement learning over time," *Neural networks for control*, pp. 67–95, 1990.

[27] D. White, D. Sofge, and L. Feldkamp, "Handbook of intelligent control: neural, fuzzy, and adaptive approaches," *IEEE Transactions on Neural Networks*, vol. 5, no. 5, pp. 852–852, 1994.

[28] B. Kim, J. Park, S. Park, and S. Kang, "Impedance learning for robotic contact tasks using natural actor-critic algorithm," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 2, pp. 433–443, 2010.

[29] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, 2011.

[30] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of dynamic systems by learning: A new control theory for servomechanism or mechatronics systems," in *Decision and Control, 1984. The 23rd IEEE Conference on*. IEEE, 1984, pp. 1064–1069.

[31] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[32] B. Siciliano and O. Khatib, *Springer handbook of robotics*. Springer Science & Business Media, 2008.

[33] A. Wahrburg, E. Morara, G. Cesari, B. Matthias, and H. Ding, "Cartesian contact force estimation for robotic manipulators using kalman filters and the generalized momentum," in *2015 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 2015, pp. 1230–1235.

[34] H. Kwakernaak and R. Sivan, *Linear optimal control systems*. Wiley-interscience New York, 1972, vol. 1.

[35] F. L. Lewis and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 1995.

[36] T. H. Lee and C. J. Harris, *Adaptive neural network control of robotic manipulators*. World Scientific, 1998, vol. 19.

[37] W. H. Young, "On classes of summable functions and their fourier series," *Proceedings of the Royal Society A Mathematical Physical and Engineering Sciences*, vol. 87, no. 594, pp. 225–229, 1912.